

# Bacterial Evolutionary Precursors of Eukaryotic Copper–Zinc Superoxide Dismutases

Gareth S.A. Wright\*

Department of Biochemistry & Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, United Kingdom

\*Corresponding author: E-mail: gareth.wright@liverpool.ac.uk.

Associate editor: Julian Echave

## Abstract

**Internalization of a bacteria by an archaeal cell expedited eukaryotic evolution. An important feature of the species that diversified into the great variety of eukaryotic life visible today was the ability to combat oxidative stress with a copper–zinc superoxide dismutase (CuZnSOD) enzyme activated by a specific, high-affinity copper chaperone. Adoption of a single protein interface that facilitates homodimerization and heterodimerization was essential; however, its evolution has been difficult to rationalize given the structural differences between bacterial and eukaryotic enzymes. In contrast, no consistent strategy for the maturation of periplasmic bacterial CuZnSODs has emerged. Here, 34 CuZnSODs are described that closely resemble the eukaryotic form but originate predominantly from aquatic bacteria. Crystal structures of a *Bacteroidetes bacterium* CuZnSOD portray both prokaryotic and eukaryotic characteristics and propose a mechanism for self-catalyzed disulfide maturation. Unification of a bacterial but eukaryotic-like CuZnSOD along with a ferredoxin-fold MXCXXC copper-binding domain within a single polypeptide created the advanced copper delivery system for CuZnSODs exemplified by the human copper chaperone for superoxide dismutase-1. The development of this system facilitated evolution of large and compartmentalized cells following endosymbiotic eukaryogenesis.**

**Key words:** metallochaperone, eukaryogenesis, oxidative stress, protein interface, SOD1.

## Introduction

Superoxide dismutases and reductases are deployed almost ubiquitously throughout nature to prevent damage to biological molecules by superoxide (Craig and Schlauch 2009; Gu and Imlay 2013; Lu et al. 2018). They employ redox-active metal cofactors to catalyze the production of molecular oxygen and hydrogen peroxide from superoxide. Iron, manganese, nickel, and copper can all serve this purpose. Iron and manganese superoxide dismutase isoenzymes evolved in bacteria prior to oxidation of the Earth's atmosphere (Dupont et al. 2010; Inupakutika et al. 2016). CuZn superoxide dismutases (CuZnSODs) evolved independently following atmospheric and oceanic oxidation as copper and zinc became bioavailable (Zerkle et al. 2005; Harel et al. 2014).

CuZnSODs from *Photobacterium leiognathi* (PhCuZnSOD) (Bourne et al. 1996), *Salmonella enterica* serovar typhimurium (Pesce et al. 2000), *Mycobacterium tuberculosis* (Spagnolo et al. 2004), and *Actinobacillus pleuropneumoniae* (Forest et al. 2000) form P-class homodimers through interface–cavity water network interactions. The P-class interface is not conserved across prokaryotic CuZnSODs, however, with P-class-like but monomeric examples found in *Escherichia coli* (Battistoni et al. 1996), *Salmonella enterica* serovar choleraesuis (Mori et al. 2008), and *Brucella abortus* (Pratt et al. 2015). These periplasmic bacterial CuZnSODs account for a small fraction of cellular superoxide dismutase activity (Steinman 1987; Benov and Fridovich 1994). In contrast, CuZnSODs are

the most abundant eukaryotic superoxide detoxifier and operate within the cytoplasm, peroxisomes, mitochondrial intermembrane space, extracellular space, and nucleus (Marklund 1982; Crapo et al. 1992; Sturtz et al. 2001). They are found in animals (McCord and Fridovich 1969), plants (Sawada et al. 1972), fungi (Gosciniak and Fridovich 1972), and protists (Ferro et al. 2015). Each conforms to a conserved, hydrophobic E-class dimer interface that incorporates four strong intersubunit hydrogen bonds (Richardson et al. 1975; Djinić et al. 1991; Lartigue et al. 2015). E-class and P-class dimerization are mechanistically distinct and the interfaces are found on opposite poles of the CuZnSOD  $\beta$ -barrel (Bourne et al. 1996).

Following translation, a sequence of post-translational modifications transform disordered, monomeric polypeptide into active CuZnSOD homodimers. Folding, zinc and copper binding, and acquisition of an intrasubunit disulfide bond can all be assisted by chaperones (Culotta et al. 1997; Battistoni et al. 1999; Osman et al. 2013; Luchinat et al. 2017; Wright 2020). For example, approximately 80% of human CuZn-superoxide dismutase-1 (SOD1) is activated by the human copper chaperone for SOD1 (hCCS) (Subramaniam et al. 2002). In addition, CCS orchestrates CuZnSOD localization to the mitochondrial intermembrane space and peroxisomes (Field et al. 2003; Kawamata and Manfredi 2008; Islinger et al. 2009). hCCS functionality rests on its ability to form E-class heterodimers with SOD1 (Schmidt et al. 1999; Banci et al.

© The Author(s) 2021. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Open Access

2012). This is mediated by an hCCS domain with 50% identity to SOD1 that conserves protein fold, zinc binding and intra-subunit disulfide bonding (Lamb et al. 2000; Sala et al. 2019). To facilitate metalation of its CuZnSOD substrate, hCCS also contains a ferredoxin-fold, MXCXXC copper-binding domain similar in structure and function to eukaryotic Atx1, prokaryotic CopZ, and related domains within the copper-transporting ATPases. Therefore, eukaryotic CuZnSOD copper chaperones likely arose through gene duplication followed by domain shuffling and sequence optimization. While intracellular E-class CuZnSOD-cognate copper chaperone pairs are very well conserved across eukaryotes they are completely absent from prokaryotes including Asgard archaea and, specifically, the Lokiarchaeota species closely related to eukaryotes (Spang et al. 2015). Thus, eukaryotic CuZnSODs and their cognate chaperone differentiated after adoption of the E-class interface but before eukaryotic radiation. Evolution of eukaryotic CuZnSODs and their chaperones has however been difficult to reconcile due to the deep division in the sequence, structure, and activation mechanisms of bacterial and eukaryotic CuZnSODs characterized thus far.

Bridging the divide between eukaryotic and prokaryotic enzymes, here I describe several classes of bacterial CuZnSODs that are more similar to the eukaryotic form than they are to previously characterized bacterial enzymes. The CuZnSOD of a *Bacteroidetes* bacterium forms its intra-subunit disulfide bond quickly and independently of chaperones. Crystal structures of this distinctly E-class protein show it includes P-class structural elements positioning several of these bacterial CuZnSODs as an evolutionary stepping-stone between E-class and P-class enzymes. The E-class interface is therefore of bacterial origin and this group of bacterial CuZnSODs are precursors of eukaryotic CuZnSOD-cognate copper chaperone pairs necessary for the development of complex eukaryotic cellular structure following endosymbiotic eukaryogenesis (Fan et al. 2020).

## Results

### Discovery of a New Group of Bacterial CuZnSODs

Bacterial and eukaryotic CuZnSOD structures available in the Protein Data Bank (PDB) have dissimilar  $\beta$ -strands 1 and 8; disulfide subloops particularly in the interface region between Gly51 and Cys57; and do not maintain a glycine residue at the human SOD1 Gly150 position (fig. 1A). In the absence of E-class dimerization, the bacterial enzymes have not been subjected to a selective pressure to adopt or maintain E-class interface residues. However, 34% of bacterial CuZnSOD sequences present in the OMA database (Altenhoff et al. 2018) retain an aspartate at the human Asp52 position within the Gly-Asp-X-Thr/Ser disulfide subloop tetrad known to report on eukaryotic CuZnSOD E-class homodimerization (fig. 1A) (Culik et al. 2018). This indicates our understanding of bacterial CuZnSOD structure may not be complete.

A protein BLAST search restricted to prokaryotic sequences using the eukaryotic CuZnSOD consensus sequence as bait yielded 34 eukaryotic-like bacterial CuZnSODs found

only in gram-negative species and distributed across 12 phyla (supplementary tables S1–S5, Supplementary Material online). Iron and manganese SOD coding sequences are often present in the same genome as P-class CuZnSODs (Benov and Fridovich 1994; Sadosky et al. 1994) and this is also observed for bacterial but eukaryotic-like CuZnSODs (Supplementary tables S1–S4, Supplementary Material online). The majority of these sequences were retrieved by environmental sampling of aquatic environments, seas, and oceans, but also water-permeating rock formations and soil. Several examples have been sequenced following laboratory culture including an example recovered from a marine hydrothermal vent (Rensink et al. 2020). Geographical sampling data indicate eukaryotic-like bacterial CuZnSODs are present globally.

The 34 bacterial CuZnSODs described in supplementary tables S1–S4, Supplementary Material online have metal-binding residues conserved in both amino acid type and order within the primary sequence compared with other eukaryotic and prokaryotic CuZnSODs confirming a monophyletic origin (fig. 1B) (Forest et al. 2000). Overall phylogenetic tree topology was stable across different amino acid substitution models, tree construction methodologies and degree of multiple sequence alignment trimming (supplementary figs. S1–S6, Supplementary Material online) with the exceptions of *Bacillus subtilis* (supplementary figs. S2, S3, and S6, Supplementary Material online) and *Megavirus chiliensis* CuZnSODs positioning (supplementary figs. S1, S3, and S4, Supplementary Material online) which have mutations that prevent copper and zinc binding, respectively. Each parameter set resulted in consensus trees with eukaryotic-like bacterial CuZnSOD groups nested within the eukaryotic CuZnSOD clade (fig. 1C and supplementary figs. S1–S6, Supplementary Material online). Indeed, figure 1C indicates bacterial eukaryotic-like CuZnSODs share a common ancestor with the eukaryotic E-class enzymes. As a whole, these CuZnSODs do not display P-class interfaces as exemplified by *PhCuZnSOD* (supplementary fig. S7, Supplementary Material online). However, conservation of eukaryotic interface surfaces, hydrogen bonding residues, and Gly150 (human numbering) indicates several of these bacterial enzymes form E-class homodimers (fig. 2). Eukaryotic-like bacterial CuZnSODs form three groups within the eukaryotic clade, each with distinct features (supplementary table S6, Supplementary Material online).

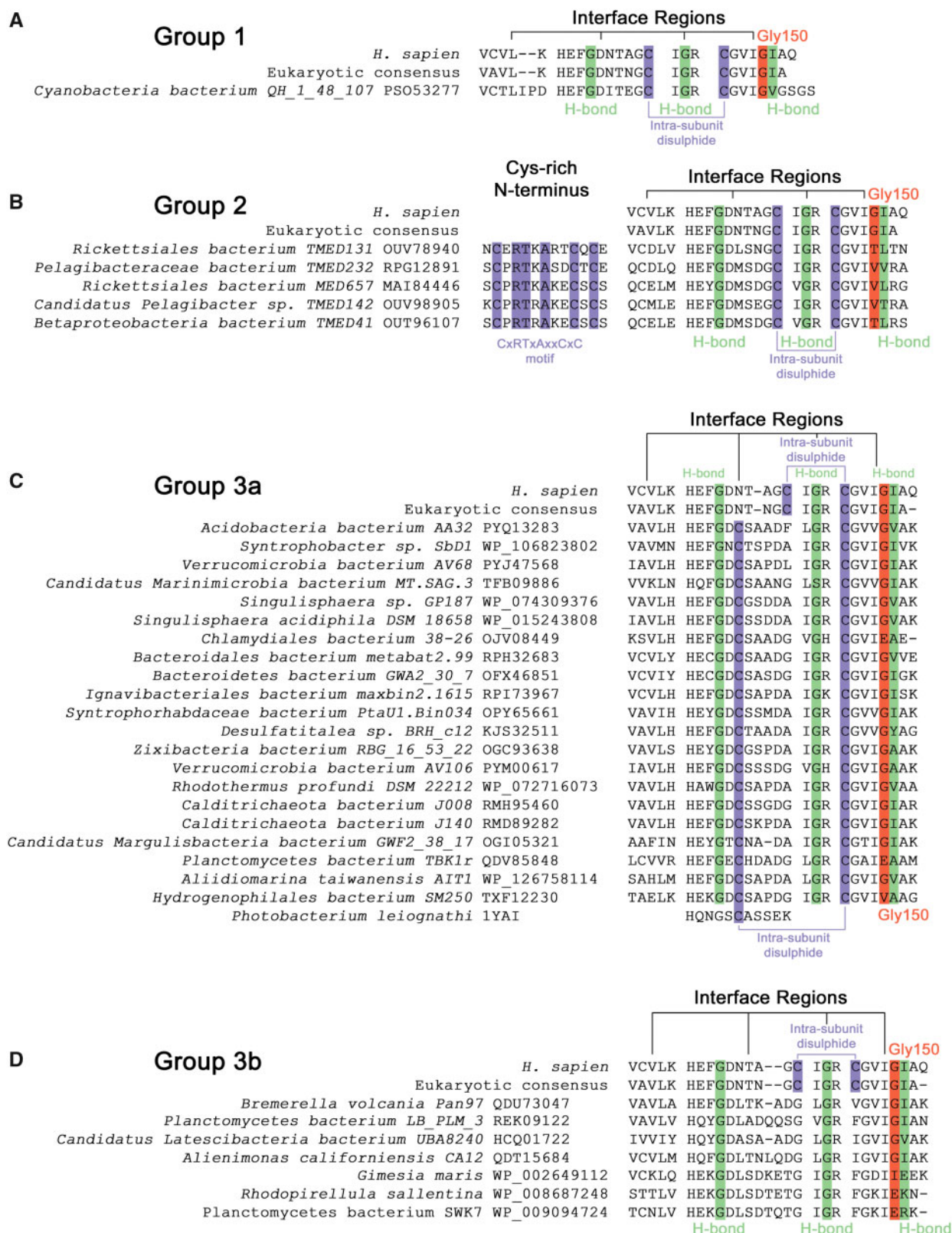
### Multiple Subgroups of Eukaryotic-Like Bacterial CuZnSODs

Group 1 comprises a single protein (supplementary table S1, Supplementary Material online) encoded within the genome of *Cyanobacterium bacterium* QH\_1\_48\_107 (PSO53277, *CbCuZnSOD*) discovered at multiple locations in the Chilean Andes (Finstad et al. 2017). No nonbacterial protein coding sequences could be found in the parent whole-genome shotgun (WGS) sequence data (supplementary dataset 1, Supplementary Material online) including coding sequences immediately adjacent to the *CbCuZnSOD* gene (supplementary table S7, Supplementary Material online).



3791





**Fig. 2.** Three groups of eukaryotic-like bacterial CuZnSODs have distinct primary structure features. (A) Group 1 *Cyanobacteria bacterium* QH\_1\_48\_107 CbCuZnSOD has E-class intrasubunit disulfide configuration and interface regions with conservative nonpolar substitutions in interface and intersubunit hydrogen bonding residues. (B) Group 2 enzymes have an N-terminal CxRTxAxxCx motif, E-class cysteine configuration, conserved hydrogen bonding residues but large and polar substitutions in place of human Gly150 and  $\beta$ -strand 1, respectively. (C) Group 3a interface residues including Gly150 (81% conserved) and intrasubunit hydrogen bonding residues are well conserved from eukaryotic CuZnSODs. Disulfide subloop cysteines involved in intrasubunit disulfide formation are conserved with respect to P-class *P. leiognathi* CuZnSOD (Bourne et al. 1996). (D) Group 3b enzymes do not have intrasubunit disulfide bonding cysteines. *G. maris*, *R. sallentina*, and *P. bacterium* SWK7 have poorly conserved E-class interface residues in contrast to other members. All include extended N-termini even after removal of signal peptide but do not contain a CXC motif unlike Group 2 enzymes. Green—interface hydrogen bonding residues, orange—Gly150 equivalent residues (human numbering) (Sala et al. 2019), purple—cysteine.

Coding sequence GC content is also well within the genomic DNA normal distribution (supplementary fig. S8, Supplementary Material online). *CbCuZnSOD* has high amino acid identity to the eukaryotic CuZnSODs consensus sequence (65.4%), but low identity to *PhCuZnSOD* (30.8%), and similar polypeptide length and predicted mass compared with intracellular, eukaryotic CuZnSODs including human SOD1 (supplementary table S1, Supplementary Material online). Like Group 2 enzymes below, its position within the eukaryotic clade (fig. 1C) indicates this bacterial CuZnSOD may result from a lateral gene transfer event from a eukaryotic source. It has E-class dimer interfaces with only conservative nonpolar for nonpolar substitutions found in  $\beta$ -strands 1 and 8 (fig. 2A). It very likely maintains the four intersubunit hydrogen bonds found in eukaryotic CuZnSODs (fig. 2A). The position of the intrasubunit disulfide bonding cysteines indicates a eukaryotic configuration linking  $\beta$ -strand 8 to the upper disulfide subloop (fig. 2A). Absence of any signal peptide to facilitate transportation through the periplasmic membrane also indicates *CbCuZnSOD* is located to the cytoplasm unlike all bacterial CuZnSODs described to date.

Group 2 *Proteobacterial* enzymes are  $54.5\% \pm 2.3$  identical to the eukaryotic consensus CuZnSOD and  $32.8\% \pm 1.9$  identical to *PhCuZnSOD* (supplementary table S2, Supplementary Material online). *Rickettsiales bacterium* TMED131 sequence data, which encodes *RbCuZnSOD* (OUV78940), was assessed for contamination by non-bacterial sources. Despite extensive bacteriophage transduction no eukaryotic or eukaryotic virus contamination was found (supplementary dataset 2 and table S8, Supplementary Material online). Group 2 CuZnSODs differ from the eukaryotic form and *CbCuZnSOD*, in that they have extended, cysteine-rich, N-termini CxRTxAxxCx motifs containing the unusual CXC triad found in the C-terminal domain of many eukaryotic CCS proteins (fig. 2B). The CXC motif binds copper (Allen et al. 2012) indicating a possible enzymatic activation function similar to the N-terminal histidine rich extensions found in some P-class CuZnSODs (Battistoni et al. 2001) or a role in disulfide formation (Banci et al. 2012). While this group largely maintains the residues involved in intersubunit hydrogen bonding, the N-terminal portion of the E-class interface is poorly conserved (fig. 2B). In addition, all of these enzymes have threonine or valine substitutions at sites analogous to human SOD1 Gly150 (fig. 2B) where even small side-chain substitutions reduce E-class homodimer affinity (Sala et al. 2019). These enzymes are therefore eukaryotic-like but predicted to be monomeric.

Group 3a is the most populous with 21 members. Mean identity to the eukaryotic consensus is  $51.3\% \pm 3.1$  whereas identity to *PhCuZnSOD* is  $33.3\% \pm 2.5$  (supplementary table S3, Supplementary Material online). *Bacteroidetes bacterium* GWA2\_30\_7 (*BbCuZnSOD*) was assessed for WGS sequence contamination from nonbacterial sources (supplementary dataset 3 and table S9, Supplementary Material online). Like P-class bacterial CuZnSODs, Group 3a leader sequences indicate they are found within the periplasmic space with one example anchored via a trans-membrane helix. Group 3a CuZnSODs as a whole have conserved E-class N- and C-

terminal interface regions but their disulfide subloop has an unusual Gly-Asp-Cys-Thr/Ser interface tetrad (fig. 2C). This cysteine positioning is ubiquitously found in P-class and P-class-like bacterial CuZnSODs but never in E-class enzymes.

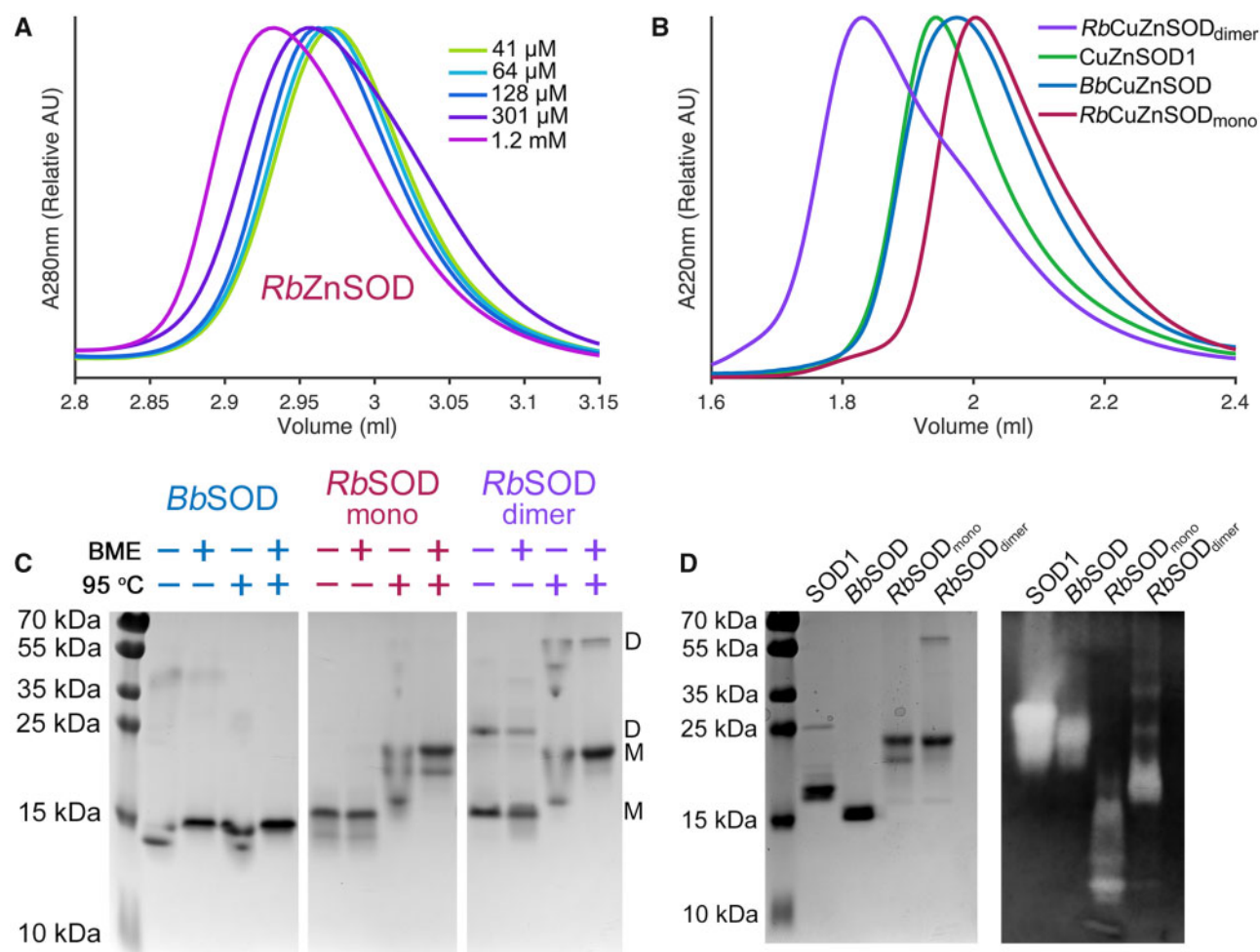
Group 3b enzymes are a heterogeneous group that belong, in the main, to the *Planctomycetes* phylum of budding bacteria (supplementary table S4, Supplementary Material online). Phylogenetic analysis indicates Group 3b is closely related to Group 3a (fig. 1C). They have mean  $54.5\% \pm 2.7$  identity to the eukaryotic CuZnSOD consensus and  $35.2\% \pm 1.6$  identity to *PhCuZnSOD*. Five are predicted to be secreted to the periplasm while two are cytoplasmic (supplementary table S4, Supplementary Material online). Their unifying characteristic is the lack of intrasubunit disulfide bonding cysteines (fig. 2D). The CuZnSOD disulfide is very well conserved across CuZnSODs with the exception of some protist enzymes (Ferro et al. 2015) but here it is replaced with a hydrophobic interaction also seen in the SOD1-like domain of *Saccharomyces cerevisiae* yCCS (fig. 2D) (Lamb et al. 1999). Four-of-seven enzymes maintain a glycine residue at the human Gly150 position, E-class interface regions and intersubunit hydrogen bonding residues. Despite this, dimer affinity is very likely to be reduced by the inability to covalently stabilize the disulfide subloop. This may be the reason three enzymes have diverged away from the canonical E-class interface with large amino acid substitutions in place of human Gly150 and Ile151 (fig. 2D).

Genome and CuZnSOD coding sequence GC analysis was performed for all 34 proteins (supplementary figs. S8–S11, Supplementary Material online). A strong correlation between genome and CuZnSOD coding sequence GC content exists (supplementary fig. S12, Supplementary Material online) and is indicative that the CuZnSOD coding sequences are habituated to their genetic, metabolic, and environmental surroundings rather than being recent additions to their respective bacterial genomes.

### Metalation, Oligomeric State and Activity of Two Eukaryotic-Like Bacterial CuZnSODs

Of the CuZnSOD sequences described above, the least likely to form an E-class homodimer interface are those with bulky side-chain amino acid substitutions in place of human SOD1 Gly150 or those with a prokaryotic-like disulfide configuration. To investigate, representative Group 2 and 3a enzymes *RbCuZnSOD* (OUV78940) and *BbCuZnSOD* (OFX46851) were purified recombinantly from *E. coli*. This yielded zinc metalated proteins with negligible copper as is the case for human SOD1 (Wright et al. 2013). Both proteins were able to bind copper provided during purification at 4.4:1 and 0.8:1 Cu:Zn ratios for *BbCuZnSOD* and *RbCuZnSOD*, respectively. Nonstoichiometric *BbCuZnSOD* copper binding is likely due to adventitious surface binding by His102-Glu104 and His102-Glu74-Asp39 sites as is the case when the protein is crystallized with zinc (7B4O). Addition of copper to the as-isolated protein does not however change the oligomeric state of the protein (supplementary fig. S13, Supplementary Material online) which has a molecular mass determined by size-exclusion chromatography coupled static light scattering





**Fig. 3.** Oligomeric state and superoxide dismutase activity of *RbCuZnSOD* and *BbCuZnSOD*. (A) SEC chromatograms showing the increase in *RbZnSOD* hydrodynamic radius as concentration increases. (B) SEC chromatograms showing distinct *RbCuZnSOD* monomers and dimers along with *BbCuZnSOD*1 and SOD1 dimers. (C) SDS-PAGE showing *RbCuZnSOD* dimers are resistant to heating and reduction in contrast to *BbCuZnSOD*. D—dimer, M—monomer, BME— $\beta$ -mercaptoethanol. (D) Reducing and denaturing SDS-PAGE (left) of *CuZnSOD*s assessed for superoxide dismutase activity by in-gel nitro tetrazolium blue staining (Beauchamp and Fridovich 1971). *RbCuZnSOD* retains activity in the monomeric state.

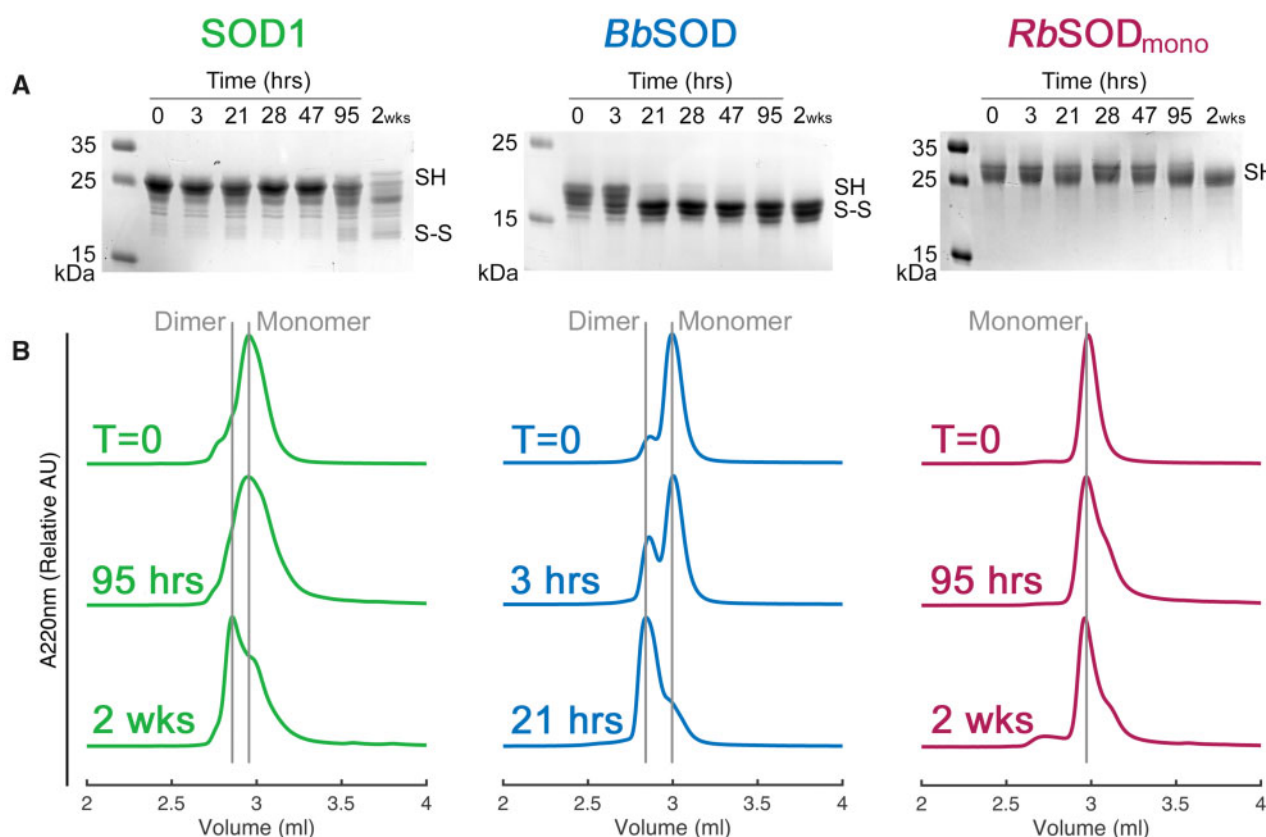
(SEC-SLS) of 30.8 kDa (polydispersity index 1.01) against a predicted dimer mass of 32.4 kDa. This indicates *BbCuZnSOD* forms a homodimer in solution.

*RbZnSOD*, in the zinc only metalated form, is predominantly monomeric and fails to fully dimerize even at millimolar concentrations (fig. 3A). On copper metalation, a species corresponding to dimeric *RbCuZnSOD* could be separated from the monomeric form (fig. 3B). *RbCuZnSOD* monomers and dimers have identical copper and zinc metalation ratios. Molecular masses calculated by SEC-SLS are 24.7 and 47.9 kDa (polydispersity indexes: 1.01 and 1.00) against expected masses from primary sequence of 19.7 and 39.5 kDa, respectively, for monomeric and dimeric species. Once formed, dimeric *RbCuZnSOD* is exceptionally stable. Dissociation into monomers on SDS-PAGE is limited even when heated to 95 °C, the presence of reductant, or a combination of both (fig. 3C). *BbCuZnSOD* and human SOD1 have similar mobility through size exclusion chromatography and native PAGE which, given their high sequence identity, similar isoelectric points (5.8 and 5.7, respectively) and solution molecular

masses, is a good indication they adopt similar quaternary structure (fig. 3B and D). *BbCuZnSOD* and *RbCuZnSOD* are active superoxide dismutases (fig. 3D). Human SOD1 is known however to lose activity when monomeric (Banci et al. 1995) but *RbCuZnSOD* monomers and dimers, which exhibit distinct mobility through native-PAGE, have equivalent dismutase activity (fig. 3D).

### Intrasubunit Disulfide Formation

Formation of the human SOD1 intrasubunit disulfide bond promotes homodimerization by creating a stable interface region where intersubunit hydrogen bonding can take place. However, the SOD1 disulfide forms slowly when functional hCCS is not present and can be reduced by cellular reductants (Banci et al. 2012; Bouldin et al. 2012). Periplasmic bacterial *CuZnSOD*s can follow a different route to maturation as exemplified by the CueP-SodC copper pathway of *S. enterica* (Osman et al. 2013) and DsbA-induced oxidative folding of *E. coli* *CuZnSOD* (Battistoni et al. 1999). Figure 4 indicates that *BbZnSOD* is able to form its intrasubunit disulfide bond



**FIG. 4.** Intrасubunit disulfide bond formation and homodimerization. (A) Reducing and denaturing SDS-PAGE showing AMS-conjugated CuZnSODs following incubation over a time course. *Bb*ZnSOD is unreactive to AMS after 21 h showing free protein thiols have been sequestered indicating intrасubunit disulfide formation. In contrast 2 weeks is necessary for human SOD1 and *Rb*ZnSOD to display significant disulfide formation. (B) SEC chromatograms showing *Bb*ZnSOD homodimerization follows the same course as intrасubunit disulfide formation. Human SOD1 does eventually form homodimers while *Rb*ZnSOD remains monomeric over the same time period.

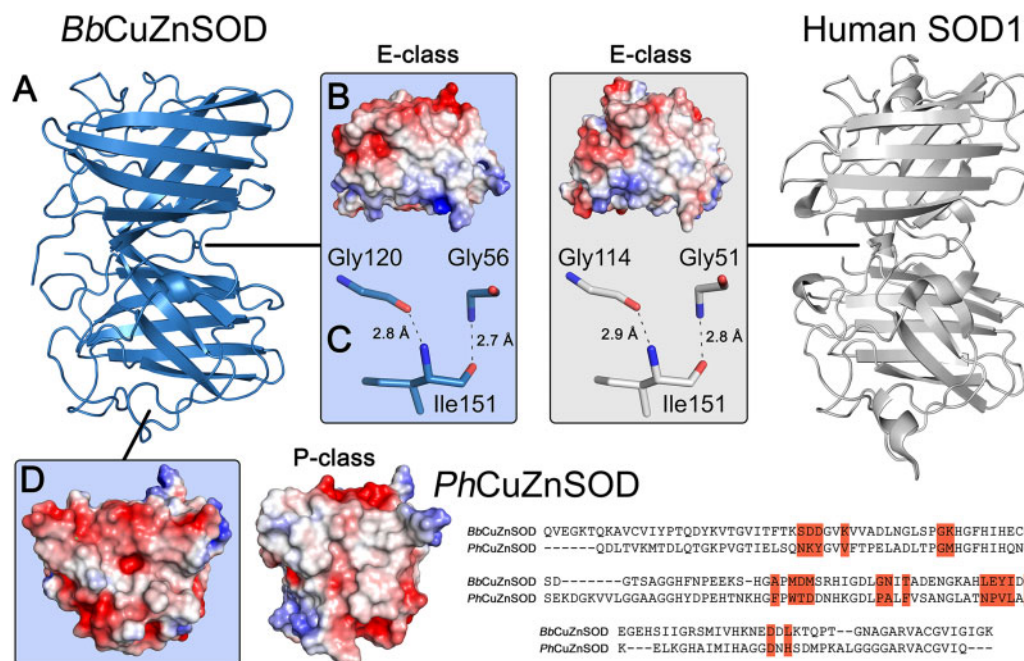
rapidly in isolation in contrast to SOD1 and *Rb*ZnSOD. As the disulfide forms, *Bb*ZnSOD transitions from monomers to homodimers indicating *Bb*CuZnSOD forms an E-class homodimer through an interface incorporating the disulfide sub-loop and that it may follow an autonomous maturation pathway.

### Structure of *Bacteroidetes bacterium* E-Class CuZnSOD

To structurally characterize this new subdivision of eukaryotic-like bacterial superoxide dismutases, *Bb*CuZnSOD was crystallized and its structure determined in zinc-zinc and copper–zinc metalation states (supplementary fig. S14, Supplementary Material online) at 1.41 Å and 2.7 Å resolution, respectively. Crystallographic statistics are presented in supplementary table S10, Supplementary Material online. *Bb*CuZnSOD has a typical eight stranded, anti-parallel, Greek-key  $\beta$ -barrel core structure with two long loops incorporating conserved metal-binding ligands and coordination geometry (Richardson et al. 1975; Parge et al. 1992). Unusually, *Bb*CuZnSOD incorporates a  $\beta$ -barrel plug residue substitution reminiscent of the ALS-related Leu106Phe mutation (supplementary fig. S15A–C, Supplementary Material online). Comparison of *Bb*CuZnSOD monomers with human SOD1, *Ph*CuZnSOD

and *E. coli* CuZnSOD monomers gives mean all-atom RMSD of 1.83 Å, 2.59 Å, and 4.22 Å, respectively. *Bb*CuZnSOD forms an E-class homodimer (fig. 5A) through a hydrophobic interface (fig. 5B) and maintains four symmetrical intersubunit hydrogen bonds typical of E-class eukaryotic CuZnSODs: Gly56-Ile151 2.7 Å and Ile151-Gly120 2.8 Å (fig. 5C). Interface hydrogen bond lengths are identical to those found in human SOD1 (fig. 5C). Strong hydrogen bonding is facilitated by conservation of Gly150 allowing close contact between opposing Gly51 and Ile151. There is no indication of a P-class interface (fig. 5D). *Bb*CuZnSOD is therefore sequentially (49.4% identity) and structurally more similar to human SOD1 than any bacterial P-class or P-class-like monomer CuZnSOD ( $32.7 \pm 2.4\%$  identity for those in the PDB), as predicted by phylogenetic analysis (fig. 1C).

The *Bb*CuZnSOD E-class dimer interface area is 747 Å<sup>2</sup> comprising N-terminus, C-terminus, and disulfide subloop regions with a PISA (Krissinel and Henrick 2007) score of 1.0 as opposed to 0.2 and 0.3 for human and bovine SOD1, respectively. The *Bb*CuZnSOD interface is larger than both human and bovine SOD1 due to aromatic stacking of opposing Tyr14 side chains in a cavity at the dimer interface (fig. 6A). In addition, each Tyr14 forms two back-bone hydrogen bonds with disulfide bonding Cys146;  $\beta$ -barrel interactions that are conserved in human SOD1. Tyrosine is not common



**FIG. 5.** Bacterial *BbCuZnSOD* dimerizes through an E-class interface. (A) Structure of homodimeric *BbCuZnSOD* showing tertiary and quaternary structure similarity to human SOD1 (2C9V). (B) Surface charge map showing the *BbCuZnSOD* interface is composed of a nonpolar patch which is conserved in human SOD1. (C) The *BbCuZnSOD* interface maintains E-class intersubunit hydrogen bonding using the same amino acid type, bond lengths, and angles as human SOD1. (D) Surface charge map showing *BbCuZnSOD* does not exhibit a nonpolar region which contributes to P-class homodimerization as exemplified by *Photobacterium leiognathi* *PhCuZnSOD* (1YAI) and has very poor conservation (9% identity) of interface residues (highlighted orange).

at this position in either eukaryotic CuZnSODs (0.5%) or the 34 bacterial enzymes presented in [supplementary tables S1–S4, Supplementary Material](#) online (9%) but occupation of this cavity with small molecules has been explored as a route to homodimer stabilization for SOD1-related ALS therapy (Ray et al. 2005; Wright et al. 2013). Immediately adjacent to Tyr14 at the loop I N-terminus, Pro15 replaces a glycine which is conserved in 80.5% of eukaryotic CuZnSODs (Gly10, human SOD1). As a result, loop I adopts a conformation shifted toward the disulfide subloop. Two extra amino acids are present in loop I allowing contact with the disulfide subloop and formation of a 2.9 Å backbone hydrogen bond between Gln17 amine and Ser61 carbonyl (fig. 6B). The disulfide subloop is itself extended by one amino acid to further accommodate this interaction. The lengths of both loops are conserved throughout the Group 3a and 3b bacterial CuZnSODs. This combination of interactions and extended loops creates a bridging network that spans the dimer interface to link intrasubunit disulfide bonds of opposing monomers (fig. 6C). This system appears to lend conformational restraint to the disulfide subloop prior to formation of the intrasubunit disulfide bond and may represent a mechanism for rapid disulfide formation (fig. 4).

#### *BbCuZnSOD* Incorporates P-Class CuZnSOD Loop Features

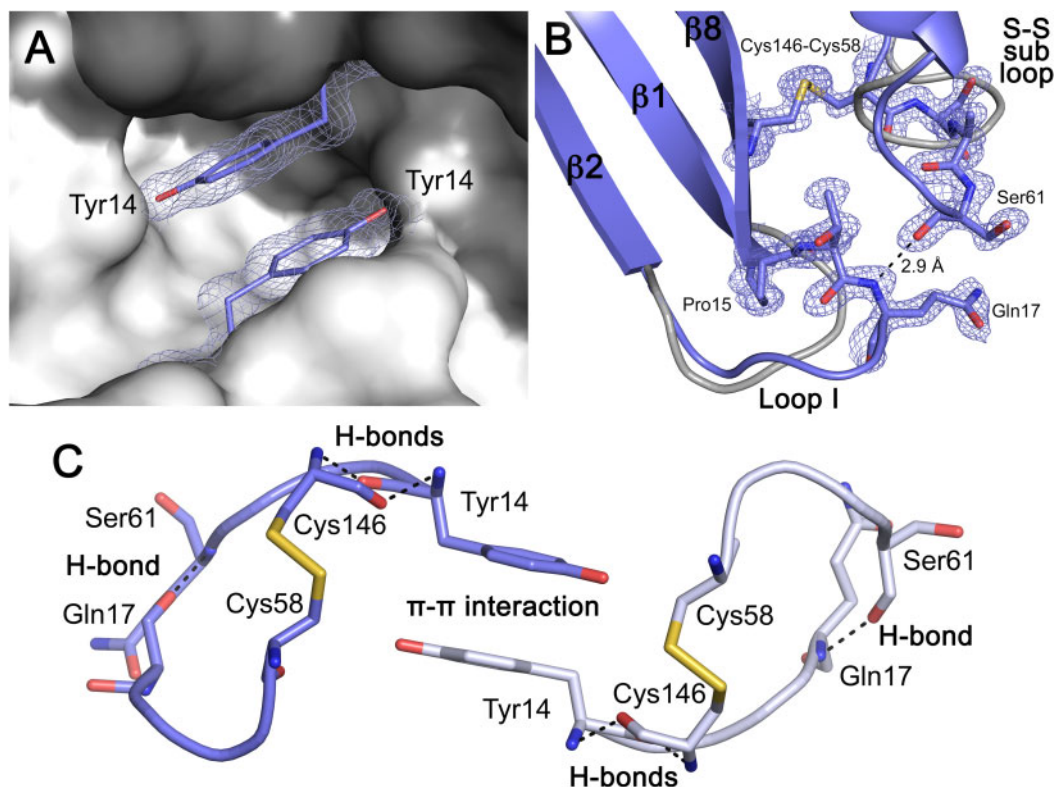
Group 3a and 3b CuZnSODs form an E-class phylogenetic group between eukaryotic and P-class CuZnSODs (fig. 1C). However, P-class commonalities are discernable in these eukaryotic-like enzymes. P-class and P-class-like CuZnSODs

have 20 or 24 amino acid disulfide subloops that extend into solvent while E-class dimers have compact 16 amino acid subloops (fig. 7A). *BbCuZnSOD*, and all Group 3a and 3b enzymes with the exception of *Candidatus margulisbacteria* CuZnSOD (OGI05321), have an intermediate 17 amino acid disulfide loop (fig. 7A). In addition, P-class enzymes form Van der Waal's and hydrophobic interactions between loop I and their disulfide suploop in a manner similar to *BbCuZnSOD* (fig. 6B). Most importantly, all Group 3a enzymes, including *BbCuZnSOD*, conform to a P-class intrasubunit disulfide configuration where the lower sub-loop links to  $\beta$ -strand 8 (figs. 2C and 7B).

A 20-amino acid electrostatic loop with a central proline is found in every P-class or P-class-like bacterial CuZnSOD characterized to date (fig. 7C). This enables a short turn around the zinc site often with no secondary structure. This feature is completely absent from all eukaryotic CuZnSODs present in the OMA database and the PDB which display 24 or 27 amino acid,  $\alpha$ -helical electrostatic loops with no central proline (fig. 7C). The electrostatic loop proline is therefore a signal characteristic of prokaryotic CuZnSODs and is conserved in *BbCuZnSOD* (Pro136) (fig. 7C) and across all Group 3a and 3b enzymes. As a result, the 18 amino acid *BbCuZnSOD* electrostatic loop more closely resembles analogous structures found in P-class and P-class-like CuZnSODs than eukaryotic CuZnSODs (fig. 7C).

A single example illustrates the chimeric nature of Group 3a proteins well. The *BbCuZnSOD* disulfide subloop and electrostatic loop have identical lengths and very similar conformations to those of a CuZnSOD-like protein of the gram-





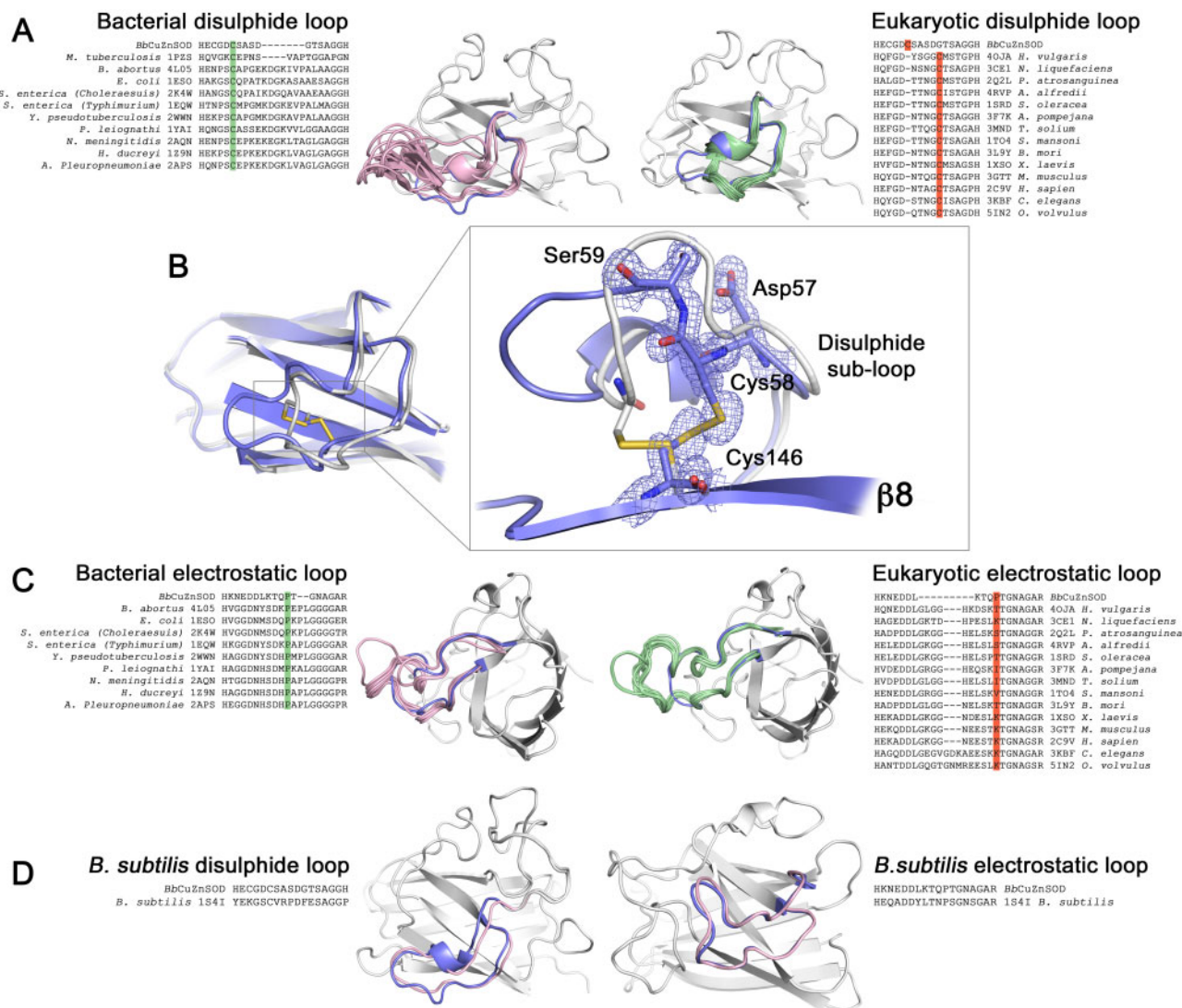
**Fig. 6.** The *BbCuZnSOD* disulfide subloop is stabilized by interactions with Try14 and loop I. (A) 2Fo-Fc electron density map contoured at  $2\sigma$  showing opposing Tyr14 residues forming side-chain hydrophobic interactions (3.4 Å) at the *BbCuZnSOD* homodimer interface aligned with the same site on the human SOD1 surface (gray and dark gray). (B) 2Fo-Fc electron density map contoured at  $2\sigma$  showing how the *BbCuZnSOD* disulfide subloop is stabilized by a hydrogen bond interaction with loop I Gln17. Both loops are longer than that found in human SOD1 (gray). (C) The *BbCuZnSOD* interface interaction network. Reciprocal and symmetric interactions described in (A) and (B) combine to stabilize disulfide subloops and the homodimer interface. Residues from opposing monomers are colored blue and gray with hydrogen bonds dashed black.

positive bacteria *Bacillus subtilis* (1S4I) (Banci et al. 2005) (fig. 7D) which is part of the P-class clade (fig. 1C). Despite this, *BbCuZnSOD* shares higher loop identity with the eukaryotic extremophile *Alvinella pompejana* CuZnSOD (3F7K) which is structurally (Shin et al. 2009) and phylogenetically E-class (figs. 1C and 7A and C).

## Discussion

Superoxide dismutases have evolved independently at least three times throughout evolution. Each new iteration was cofactored by different redox metals as environmental conditions dictated available resources. CuZnSODs are present in all three domains of life: eukaryotes (McCord and Fridovich 1969; Gosciniak and Fridovich 1972; Sawada et al. 1972; Ferro et al. 2015), bacteria (Battistoni et al. 1996; Bourne et al. 1996; Forest et al. 2000; Pesce et al. 2000; Spagnolo et al. 2004; Mori et al. 2008; Pratt et al. 2015), and archaea (supplementary table S11, Supplementary Material online). All have the same metal-coordinating ligands in the same order within the primary structure indicating a monophyletic origin with metal binding and acquisition of activity predating evolution of eukaryotes (Forest et al. 2000). Despite a single origin, bacterial and eukaryotic CuZnSODs characterized to date differ

considerably with respect to sequence, loop structures, and dimerization interface (Parge et al. 1992; Bourne et al. 1996). The Group 3 CuZnSODs described here and exemplified by *BbCuZnSOD* represent an evolutionary intermediary between E-class and P-class enzymes with high eukaryotic sequence identity and E-class homodimerization but P-class functional loops and disulfide configuration. The global distribution of E-class bacterial CuZnSODs, their occurrence in many different phyla and adaptation of coding sequence GC content to their parent genomes indicate they are very likely to have been present at the time of eukaryogenesis. Despite exceptions within the protists and nematodes, which have lost antioxidant enzymes including CuZnSODs over time, E-class CuZnSODs are extremely well conserved across eukaryotes indicating they were present in the last common eukaryotic ancestor (McCord and Fridovich 1969; Gosciniak and Fridovich 1972; Sawada et al. 1972; Ferro et al. 2015). The absence of archaeal E-class CuZnSODs (supplementary table S11, Supplementary Material online) and the well described parallels between eukaryotic mitochondria and proteobacteria, particularly Rickettsiales species (Fan et al. 2020), indicate that an ancestral bacterial but eukaryotic-like CuZnSOD arrived in the proto-eukaryotic cell through endosymbiosis. That early eukaryotic



**FIG. 7.** Eukaryotic-like *BbCuZnSOD*  $\beta$ -barrel loops display similarities to P-class CuZnSODs. (A) Sequence and structural comparison of the *BbCuZnSOD* disulphide subloop with analogous P-class and E-class CuZnSOD structures shows *BbCuZnSOD* loops are of intermediate amino acid length and conformational extension. Structural features are highlighted pink—bacterial, green—eukaryotic, and blue—*BbCuZnSOD* with all structures aligned to the *BbCuZnSOD*  $\beta$ -barrel (gray). Sequence data show *BbCuZnSOD* has a bacterial disulfide subloop cysteine configuration (green) as opposed to eukaryotic (orange). (B) 2Fo-Fc electron density map contoured at  $2\sigma$  showing *BbCuZnSOD* has a characteristically prokaryotic intrasubunit disulfide bond configuration (blue) in comparison with human SOD1 (gray). (C) Sequence and structural comparison of the *BbCuZnSOD* electrostatic loop with bacterial and eukaryotic structures. The *BbCuZnSOD* has length and conformation similar to bacterial orthologs with a characteristic and highly conserved central proline (green) that is absent in eukaryotic enzymes (orange). Structural features are colored as in (A). (D) Comparison of *BbCuZnSOD* loop sequence and conformation with that of *B. subtilis* (1541). *BbCuZnSOD* loop identity to *B. subtilis* is 47% and 56% in comparison with *Alvinella pompejana* 59% and 61% for disulfide subloop and electrostatic loops, respectively (gaps excluded). Structural features are colored as in (A).

CuZnSOD could be represented today by Group 1 or proteobacterial Group 2 enzymes with E-class disulfide and loop structures or Group 3a enzymes with remnant P-class features.

E-class CuZnSOD dimers have higher metal-binding affinity but lower activity than P-class dimers (Cabbianelli et al. 2004). This facilitates copper retention by intracellular enzymes but the E-class interface also allows engagement with a single, specific, and high-affinity cognate molecular, disulfide, and copper chaperone. CuZnSODs that have sufficiently high copper affinity to receive their cofactor from this metallochaperone distribution system allow maturation

distant from the cytoplasmic membrane without reliance on metal ion diffusion (Pope et al. 2013) and within membrane-bound cellular compartments. The bacterial E-class CuZnSODs described here along with the copper-binding MXCXXC motif proteins typified by CopZ and found in all the bacterial phyla represented in [supplementary table S5, Supplementary Material](#) online are protein-domain building blocks requiring only duplication and shuffling to create a fully functional intracellular copper chaperone delivery system CuZnSODs resembling human and yeast CCS proteins. This system was in place before diversification of the last eukaryotic common ancestor where it facilitated the

existence of large and organizationally complex cells within multicellular organisms.

## Materials and Methods

### Sequence Analysis

The eukaryotic CuZnSOD consensus sequence was determined by alignment of eukaryotic CuZnSOD orthologs downloaded from the OMA database (Altenhoff et al. 2018) in Jalview (Waterhouse et al. 2009). Putative eukaryotic-like bacterial CuZnSOD protein sequences were retrieved using the protein–protein BLAST (BLASTP) web interface and querying the eukaryotic CuZnSOD consensus sequence against the NCBI nonredundant (nr) database restricted to bacteria (taxid : 2) with default settings. Sequences were assessed for loosely conserved CuZnSOD disulfide sub-loop Gly-Asp-Gly-Cys and  $\beta$ -strand 1 and 2 sequences. Putative eukaryotic-like bacterial CuZnSOD sequences were then used to query the nr database with server BLASTP restricted to nonbacterial sequences. This yielded several protein sequences that had very high identity to a eukaryotic CuZnSOD including an annotated *Acinetobacter baumannii* protein (WP\_071414557) with 97% identity to Scotts pine CuZnSOD (P24669). All were removed from subsequent analysis as they were likely due to sample contamination. Each putative bacterial CuZnSOD was also pair-wise aligned with human SOD1 and *PhCuZnSOD* after removal of any signal peptide indicated by SignalP-5.0 (Almagro Armenteros et al. 2019). Transmembrane regions were predicted with TMHMM 2.0 (Möller et al. 2001). Identities are recorded in [supplementary tables S1–S4, Supplementary Material](#) online. Bacterial CuZnSODs were aligned with Muscle (Edgar 2004) and assigned to different groups based primary sequence motifs including disulfide cysteine positioning, signal peptide, and cysteine-rich N-terminus.

*Cyanobacterium bacterium* QH\_1\_48\_107, *Rickettsiales bacterium* TMED131, and *Bacteroidetes bacterium* GWA2\_30\_7 WGS sequencing data (PXPJ00000000, NHGG00000000, and MENC00000000, respectively) were separated into individual contigs and used as queries against the NCBI reference sequence (RefSeq) database (O’Leary et al. 2016) restricted to nonbacterial sequences using standalone BLASTX 2.10.0+ (Altschul et al. 1997). For each contig, a protein translation with the highest BLAST score to a nonbacterial protein sequence was used to query the RefSeq database restricted to bacterial proteins using server BLASTP. Results were filtered to remove proteins from organisms sequenced as part of the same study. The difference between BLAST scores for bacterial and nonbacterial sequences were then used to assess the likelihood genomic assemblies contained nonbacterial material. The top hit from each contig is recorded in [supplementary datasets 1–3, Supplementary Material](#) online. Translations of coding sequences immediately adjacent to *CbCuZnSOD*, *RbCuZnSOD*, and *BbCuZnSOD* coding sequences were also used to interrogate the RefSeq database using BLASTP restricted to bacterial or nonbacterial sequences ([supplementary tables S7–S9, Supplementary Material](#) online,

respectively). Genome GC content was analyzed using the Matlab `ntdensity` function.

### Phylogenetic Tree Building

To limit the effects of genetic saturation, protein sequences of every CuZnSOD found in the PDB and those listed in [supplementary tables S1–S4, Supplementary Material](#) online were aligned with Muscle (Edgar 2004) following removal of any signal peptide. The full alignment of 63 taxons and 332 characters was trimmed with Gblocks (Talavera and Castresana 2007) to remove divergent, misaligned, or gapped positions using both strict and loose parameters yielding alignments of 34 and 96 characters, respectively. An alignment trimmed to incorporate only the CuZnSOD  $\beta$ -barrel with internal loops and minimize gaps was created manually in Jalview (Waterhouse et al. 2009). Selection of best fit amino acid substitution rate models was performed for each alignment using the Akaike information criterion implemented in PhyML 3.0 (Guindon et al. 2010) and MEGA 10.8 (Stecher et al. 2020). Dayhoff (Dayhoff et al. 1978), Jones–Taylor–Thornton (JTT) (Jones et al. 1992) and Whelan–Goldman (WAG) (Whelan and Goldman 2001) fixed rate models were predicted with roughly equal validity. Phylogenetic analysis was performed by Bayesian inference with MrBayes 3.2 (Ronquist et al. 2012) with a gamma-distributed rate variation across all sites (G), a proportion of invariant sites combined with a gamma distribution for other sites (invG), and an invariant distribution across all sites (inv) for all alignments described above. Phylogenetic trees were also constructed using Le-Gascuel (LG), Blosum, and rtRev fixed rate models and GTR and Equalin variable rate models. Tree building was performed for 38 different parameter or alignment sets until the average standard deviation of split frequencies between two independent runs of four Markov Chain Monte Carlo (MCMC) chains was less than 0.005. The first 25% of sampled trees were discarded during analysis. P-class/P-class-like sequences and E-class/E-class-like sequences were separated into distinct clades for every tree constructed with 100% probability. The maximum likelihood method was also used to independently verify phylogenetic tree form using PhyML 3.0 (Guindon et al. 2010) using the WAG model. Unrooted trees were drawn with FigTree 1.4.4.

### Protein Expression and Purification

*Rickettsiales bacterium* TMED131 *RbCuZnSOD* and *Bacteroidetes bacterium* GWA2\_30\_7 *BbCuZnSOD* coding DNA was synthesized by Twist Bioscience along with an N-terminal hexa-histidine and SUMO tags and ligated into pET21 expression vector. Both proteins were expressed in BL21 (DE3) *E. coli* in LB media at 18 °C for 16 h with shaking at 180 rpm following induction with 0.4 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside and 200  $\mu$ M ZnCl<sub>2</sub>. Cells were lysed by sonication in 20 mM tris(hydroxymethyl)aminomethane-HCl pH 7.4, 500 mM NaCl, 20 mM imidazole, bound to Nickel-nitrilotriacetic acid resin and eluted with the same buffer including 500 mM imidazole. The His-SUMO tag was removed with SUMO protease in dialysis at 4 °C for 16 h against 20 mM tris(hydroxymethyl)aminomethane-HCl pH 7.4,



150 mM NaCl to yield the native protein sequence with no additional N-terminal amino acids. The protein was then concentrated and separated from remaining impurities on a Superdex S200 16 600 SEC column. Protein concentration was quantified by Bradford assay. Copper metalated protein was produced by dialysis against 20 mM tris(hydroxymethyl)aminomethane-HCl pH 7.4, 150 mM NaCl, 2 mM CuCl<sub>2</sub> followed by 3-fold dialysis overnight against the same buffer without copper prior to preparative SEC. Samples for inductively coupled plasma metal analysis were prepared by incubation in 70% nitric acid for 24 h at room temperature before to dilution to 4.9% nitric acid.

### Analytical SEC and Molecular Mass Determination

An Agilent BioSEC Advance 300 Å, 4.6 × 300 mm or GE Superdex 200 5 15 column was used for analytical SEC and SEC-SLS in 20 mM tris(hydroxymethyl)aminomethane-HCl pH 7.4, 150 mM NaCl buffer. Both were performed on an Agilent 1260 Infinity Multi-Detector System using refractive index to calculate masses.

### In-Gel Superoxide Dismutase Activity

CuZnSOD activity was tested in gel according to an established protocol (Beauchamp and Fridovich 1971) where superoxide dismutase activity inhibits formation of blue formazan from nitroterazolum blue by superoxide generated by the photoreduction of riboflavin thereby creating colorless bands on a dark blue gel background. Briefly, 750 ng of protein was loaded on a 12% acrylamide nonreducing native PAGE gel and run at 100 V until the dye front left the gel. The gel was then washed in 50 mM K<sub>2</sub>HPO<sub>4</sub> pH 7.5 for 10 min; 480 μM nitroterazolum blue chloride, 50 mM K<sub>2</sub>HPO<sub>4</sub> pH 7.5 for 20 min, 30 μM riboflavin, 0.02% N,N,N',N'-tetramethylethylenediamine, 50 mM K<sub>2</sub>HPO<sub>4</sub> pH 7.5 for 20 min. The gel was then exposed to white light until banding became clear. About 1.25 μg of protein from the same samples, with the addition of 2-mercaptoethanol and heating at 95 °C for 5 min, was also separated by reducing and denaturing SDS-PAGE and stained by Coomassie blue.

### Disulfide Formation Assay

Zinc-bound, copper-free, disulfide-reduced proteins were produced by incubation of as-isolated superoxide dismutases with 40 mM dithiothreitol at 4 °C overnight. Reductant was removed by desalting into nitrogen-purged 20 mM tris(hydroxymethyl)aminomethane-HCl pH 7.4, 150 mM NaCl with PD Mini-tap G25 columns pre-equilibrated with buffer. Eluted proteins were diluted to 20 μM with nitrogen-purged buffer and incubated at room temperature for the experiment time course. Samples were taken periodically and analyzed by analytical SEC, as described above, or incubated with 400 μM 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid (AMS) for 90 min at 37 °C prior to separation by denaturing and reducing 15% acrylamide SDS-PAGE.

### Crystallization

As isolated BbZnSOD, lacking catalytic copper, was crystallized at 16 mg/ml using the sitting drop vapor diffusion method in 100 mM tris(hydroxymethyl)aminomethane-HCl pH 8, 20% polyethylene glycol 6000, 2 mM ZnCl<sub>2</sub> at 22 °C for approximately 4 weeks. Crystals were cryoprotected in 100 mM tris(hydroxymethyl)aminomethane-HCl pH 8, 20% polyethylene glycol 6000, 20% glycerol, and flash frozen in liquid nitrogen. BbCuZnSOD was crystallized as above but in Morpheus (Molecular Dimensions) screen condition 1-44 comprising: 0.12 M alcohol mix containing 1,6-hexanediol, 1-butanol 1,2-propanediol, 2-propanol, 4-butanediol, 1,3-propanediol; 0.1 M buffer mix containing sodium 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, 3-(N-morpholino)propane sulfonic acid pH 7.5; 37.5% precipitant mix containing 2-Methyl-2,4-pentanediol, polyethylene glycol 1000, polyethylene glycol 3350, and flash frozen in the drop solution.

### Crystallographic Data Collection and Structure Solution

Diffraction data was collected at Synchrotron Soleil on beam-line Proxima 1 at wavelength 0.98, 1.22, and 1.33 Å where appropriate. Data were indexed and integrated with XDS (Kabsch 2010) or iMosflm (Battye et al. 2011); scaled with Aimless within CCP4 (Winn et al. 2011); solved by molecular replacement using the human SOD1 structure (2C9V) as a search model, or the zinc-metalated BbZnSOD structure for the Cu-Zn metalated variant, with Molrep (Vagin and Teplyakov 1997); rebuilt with ARP/wARP (Langer et al. 2008, p. 7); and refined with Refmac (Murshudov et al. 2011, p. 5) and Coot (Emsley et al. 2010). Structures were validated with Molprobtity and the PDB validation tool and deposited in the PDB with accession codes 7B4O and 7B4P. Electron density maps were created with Phenix Maps (Liebschner et al. 2019) and Pymol.

### Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

### Acknowledgments

Thanks to the staff of beam-line Proxima 1 at Synchrotron Soleil, France. Thanks also to Dr Svetlana Antonyuk, Prof. Alistair Darby, and Dr Jonathan Bollback for useful discussions and advice. This work was supported by the Motor Neurone Disease Association (Wright/Oct18/969-799).

### Author Contributions

Conceptualization, methodology, investigation, writing, and editing: G.S.A.W.

### Data Availability

Crystal structures are deposited in the Protein Data Bank with accession codes 7B4O and 7B4P. All data and materials are available, without restriction, upon request.

## References

- Allen S, Badarau A, Dennison C. 2012. Cu(I) affinities of the domain 1 and 3 sites in the human metallochaperone for Cu,Zn-superoxide dismutase. *Biochemistry* 51(7):1439–1448.
- Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*. 37(4):420–423.
- Altenhoff AM, Glover NM, Train C-M, Kaleb K, Warwick Vesztrocy A, Dylus D, de Farias TM, Zile K, Stevenson C, Long J, et al. 2018. The OMA orthology database in 2018: retrieving evolutionary relationships among all domains of life through richer web and programmatic interfaces. *Nucleic Acids Res*. 46(D1):D477–D485.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 25(17):3389–3402.
- Banci L, Bertini I, Calderone V, Cramaro F, Del Conte R, Fantoni A, Mangani S, Quattrone A, Viezzoli MS. 2005. A prokaryotic superoxide dismutase paralog lacking two Cu ligands: from largely unstructured in solution to ordered in the crystal. *Proc Natl Acad Sci U S A*. 102(21):7541–7546.
- Banci L, Bertini I, Cantini F, Kozyreva T, Massagni C, Palumaa P, Rubino JT, Zovo K. 2012. Human superoxide dismutase 1 (hSOD1) maturation through interaction with human copper chaperone for SOD1 (hCCS). *Proc Natl Acad Sci U S A*. 109(34):13555–13560.
- Banci L, Bertini I, Chiu CY, Mullenbach GT, Viezzoli MS. 1995. Synthesis and characterization of a monomeric mutant Cu/Zn superoxide dismutase with partially reconstituted enzymic activity. *Eur J Biochem*. 234(3):855–860.
- Battistoni A, Folcarelli S, Gabbianelli R, Capo C, Rotilio G. 1996. The Cu,Zn superoxide dismutase from *Escherichia coli* retains monomeric structure at high protein concentration. Evidence for altered subunit interaction in all the bacteriocupreins. *Biochem J*. 320(3):713–716.
- Battistoni A, Mazzetti AP, Rotilio G. 1999. In vivo formation of Cu,Zn superoxide dismutase disulfide bond in *Escherichia coli*. *FEBS Lett*. 443(3):313–316.
- Battistoni A, Pacello F, Mazzetti AP, Capo C, Kroll JS, Langford PR, Sansone A, Donnarumma G, Valenti P, Rotilio G. 2001. A histidine-rich metal binding domain at the N terminus of Cu,Zn-superoxide dismutases from pathogenic bacteria: a novel strategy for metal chaperoning. *J Biol Chem*. 276(32):30315–30325.
- Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr*. 67(Pt 4):271–281.
- Beauchamp C, Fridovich I. 1971. Superoxide dismutase: improved assays and an assay applicable to acrylamide gels. *Anal. Biochem*. 44(1):276–287.
- Benov LT, Fridovich I. 1994. *Escherichia coli* expresses a copper- and zinc-containing superoxide dismutase. *J Biol Chem*. 269(41):25310–25314.
- Bouldin SD, Darch MA, Hart PJ, Outten CE. 2012. Redox properties of the disulfide bond of human Cu,Zn superoxide dismutase and the effects of human glutaredoxin 1. *Biochem J*. 446(1):59–67.
- Bourne Y, Redford SM, Steinman HM, Lepock JR, Tainer JA, Getzoff ED. 1996. Novel dimeric interface and electrostatic recognition in bacterial Cu,Zn superoxide dismutase. *Proc Natl Acad Sci U S A*. 93(23):12774–12779.
- Craig M, Schlauch JM. 2009. Phagocytic superoxide specifically damages an extracytoplasmic target to inhibit or kill *Salmonella*. *PLoS One*. 4(3):e4975.
- Crapo JD, Oury T, Rabouille C, Slot JW, Chang LY. 1992. Copper, zinc superoxide dismutase is primarily a cytosolic protein in human cells. *Proc Natl Acad Sci U S A*. 89(21):10405–10409.
- Culik RM, Sekhar A, Nagesh J, Deol H, Rumfeldt JAO, Meiering EM, Kay LE. 2018. Effects of maturation on the conformational free-energy landscape of SOD1. *Proc Natl Acad Sci U S A*. 115(11):E2546–E2555.
- Culotta VC, Klomp LW, Strain J, Casareno RL, Krems B, Gitlin JD. 1997. The copper chaperone for superoxide dismutase. *J Biol Chem*. 272(38):23469–23472.
- Dayhoff MO, Schwartz RM, Orcutt BC. 1978. A model of evolutionary change in proteins. In: Dayhoff MO, editor. Atlas of protein sequence and structure. Vol. 5. Washington (DC): National Biomedical Research Foundation. p. 345–352.
- Djinović K, Gatti G, Coda A, Antolini L, Pelosi G, Desideri A, Falconi M, Marmocchi F, Rotilio G, Bolognesi M. 1991. Structure solution and molecular dynamics refinement of the yeast Cu,Zn enzyme superoxide dismutase. *Acta Crystallogr B Struct Sci*. 47(6):918–927.
- Dupont CL, Butcher A, Valas RE, Bourne PE, Caetano-Anollés G. 2010. History of biological metal utilization inferred through phylogenomic analysis of protein structures. *Proc Natl Acad Sci U S A*. 107(23):10567–10572.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32(5):1792–1797.
- Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 66(Pt 4):486–501.
- Fan L, Wu D, Goremykin V, Xiao J, Xu Y, Garg S, Zhang C, Martin WF, Zhu R. 2020. Phylogenetic analyses with systematic taxon sampling show that mitochondria branch within Alphaproteobacteria. *Nat Ecol Evol*. 4(9):1213–1219.
- Ferro D, Bakiu R, De Pittà C, Boldrin F, Cattalini F, Pucciarelli S, Miceli C, Santovito G. 2015. Cu,Zn superoxide dismutases from *Tetrahymena thermophila*: molecular evolution and gene expression of the first line of antioxidant defenses. *Protist* 166(1):131–145.
- Field LS, Furukawa Y, O'Halloran TV, Culotta VC. 2003. Factors controlling the uptake of yeast copper/zinc superoxide dismutase into mitochondria. *J Biol Chem*. 278(30):28052–28059.
- Finstad KM, Probst AJ, Thomas BC, Andersen GL, Demergasso C, Echeverría A, Amundson RG, Banfield JF. 2017. Microbial community structure and the persistence of cyanobacterial populations in salt crusts of the hyperarid atacama desert from genome-resolved metagenomics. *Front Microbiol*. 8:1435.
- Forest KT, Langford PR, Kroll JS, Getzoff ED. 2000. Cu, Zn superoxide dismutase structure from a microbial pathogen establishes a class with a conserved dimer interface. *J Mol Biol*. 296(1):145–153.
- Gabbianelli R, D'Orazio M, Pacello F, O'Neill P, Nicolini L, Rotilio G, Battistoni A. 2004. Distinctive functional features in prokaryotic and eukaryotic Cu,Zn superoxide dismutases. *Biol Chem*. 385(8):749–754.
- Goscini SA, Fridovich I. 1972. The purification and properties of superoxide dismutase from *Saccharomyces cerevisiae*. *Biochim Biophys Acta*. 289(2):276–283.
- Gu M, Imlay JA. 2013. Superoxide poisons mononuclear iron enzymes by causing mismetallation. *Mol Microbiol*. 89(1):123–134.
- Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 59(3):307–321.
- Harel A, Bromberg Y, Falkowski PG, Bhattacharya D. 2014. Evolutionary history of redox metal-binding domains across the tree of life. *Proc Natl Acad Sci U S A*. 111(19):7042–7047.
- Inupakutika MA, Sengupta S, Devireddy AR, Azad RK, Mittler R. 2016. The evolution of reactive oxygen species metabolism. *J Exp Bot*. 67(21):5933–5943.
- Islinger M, Li KW, Seitz J, Völkl A, Lüers GH. 2009. Hitchhiking of Cu/Zn superoxide dismutase to peroxisomes—evidence for a natural piggyback import mechanism in mammals. *Traffic* 10(11):1711–1721.
- Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*. 8(3):275–282.
- Kabsch W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr*. 66(Pt 2):125–132.

- Kawamata H, Manfredi G. 2008. Different regulation of wild-type and mutant Cu,Zn superoxide dismutase localization in mammalian mitochondria. *Hum Mol Genet.* 17(21):3303–3317.
- Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol.* 372(3):774–797.
- Lamb AL, Wernimont AK, Pufahl RA, Culotta VC, O'Halloran TV, Rosenzweig AC. 1999. Crystal structure of the copper chaperone for superoxide dismutase. *Nat Struct Biol.* 6(8):724–729.
- Lamb AL, Wernimont AK, Pufahl RA, O'Halloran TV, Rosenzweig AC. 2000. Crystal structure of the second domain of the human copper chaperone for superoxide dismutase. *Biochemistry* 39(7):1589–1595.
- Langer G, Cohen SX, Lamzin VS, Perrakis A. 2008. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc.* 3(7):1171–1179.
- Lartigue A, Burlat B, Coutard B, Chaspoul F, Claverie J-M, Abergel C. 2015. The megavirus chilensis Cu,Zn-superoxide dismutase: the first viral structure of a typical cellular copper chaperone-independent hyperstable dimeric enzyme. *J Virol.* 89(1):824–832.
- Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, et al. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol.* 75(Pt 10):861–877.
- Lu Z, Sethu R, Imlay JA. 2018. Endogenous superoxide is a key effector of the oxygen sensitivity of a model obligate anaerobe. *Proc Natl Acad Sci U S A.* 115(14):E3266–E3275.
- Luchinat E, Barbieri L, Banci L. 2017. A molecular chaperone activity of CCS restores the maturation of SOD1 fALS mutants. *Sci Rep.* 7(1):17433.
- Marklund SL. 1982. Human copper-containing superoxide dismutase of high molecular weight. *Proc Natl Acad Sci U S A.* 79(24):7634–7638.
- McCord JM, Fridovich I. 1969. Superoxide dismutase. An enzymic function for erythrocyte (hemocuprein). *J Biol Chem.* 244(22):6049–6055.
- Möller S, Croning MD, Apweiler R. 2001. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17(7):646–653.
- Mori M, Jiménez B, Piccoli M, Battistoni A, Sette M. 2008. The solution structure of the monomeric copper, zinc superoxide dismutase from *Salmonella enterica*: structural insights to understand the evolution toward the dimeric structure. *Biochemistry* 47(49):12954–12963.
- Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr.* 67(Pt 4):355–367.
- O'Leary NA, Wright MW, Brister JR, Ciufio S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, et al. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44(D1):D733–D745.
- Osman D, Patterson CJ, Bailey K, Fisher K, Robinson NJ, Rigby SE, Cavet JS. 2013. The copper supply pathway to a *Salmonella* Cu,Zn-superoxide dismutase (SodCII) involves P(1B)-type ATPase copper efflux and periplasmic CueP. *Mol Microbiol.* 87(3):466–477.
- Parge HE, Hallewell RA, Tainer JA. 1992. Atomic structures of wild-type and thermostable mutant recombinant human Cu,Zn superoxide dismutase. *Proc Natl Acad Sci U S A.* 89(13):6109–6113.
- Pesce A, Battistoni A, Stroppolo ME, Polizio F, Nardini M, Kroll JS, Langford PR, O'Neill P, Sette M, Desideri A, et al. 2000. Functional and crystallographic characterization of *Salmonella typhimurium* Cu,Zn superoxide dismutase coded by the sodCI virulence gene. *J Mol Biol.* 302(2):465–478.
- Pope CR, De Feo CJ, Unger VM. 2013. Cellular distribution of copper to superoxide dismutase involves scaffolding by membranes. *Proc Natl Acad Sci U S A.* 110(51):20491–20496.
- Pratt AJ, DiDonato M, Shin DS, Cabelli DE, Bruns CK, Belzer CA, Gorringer AR, Langford PR, Tabatabai LB, Kroll JS, et al. 2015. Structural, functional, and immunogenic insights on Cu,Zn superoxide dismutase pathogenic virulence factors from *Neisseria meningitidis* and *Brucella abortus*. *J Bacteriol.* 197(24):3834–3847.
- Ray SS, Nowak RJ, Brown RH, Lansbury PT. 2005. Small-molecule-mediated stabilization of familial amyotrophic lateral sclerosis-linked superoxide dismutase mutants against unfolding and aggregation. *Proc Natl Acad Sci U S A.* 102(10):3639–3644.
- Rensink S, Wiegand S, Kallscheuer N, Rast P, Peeters SH, Heuer A, Boedeker C, Jetten MSM, Rohde M, Jogler M, et al. 2020. Description of the novel planctomycetal genus *Bremerella*, containing *Bremerella volcania* sp. nov., isolated from an active volcanic site, and reclassification of *Blastopirellula crenea* as *Bremerella crenea* comb. nov. *Antonie van Leeuwenhoek* 113(12):1823–1837.
- Richardson J, Thomas KA, Rubin BH, Richardson DC. 1975. Crystal structure of bovine Cu,Zn superoxide dismutase at 3 Å resolution: chain tracing and metal ligands. *Proc Natl Acad Sci U S A.* 72(4):1349–1353.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 61(3):539–542.
- Sadosky AB, Wilson JW, Steinman HM, Shuman HA. 1994. The iron superoxide dismutase of *Legionella pneumophila* is essential for viability. *J Bacteriol.* 176(12):3790–3799.
- Sala FA, Wright GSA, Antonyuk SV, Garratt RC, Hasnain SS. 2019. Molecular recognition and maturation of SOD1 by its evolutionarily destabilised cognate chaperone hCCS. *PLoS Biol.* 17(2):e3000141.
- Sawada Y, Oyama T, Yamazaki I. 1972. Preparation and physicochemical properties of green pea superoxide dismutase. *Biochim Biophys Acta.* 268(2):305–312.
- Schmidt PJ, Rae TD, Pufahl RA, Hamma T, Strain J, O'Halloran TV, Culotta VC. 1999. Multiple protein domains contribute to the action of the copper chaperone for superoxide dismutase. *J Biol Chem.* 274(34):23719–23725.
- Shin DS, DiDonato M, Barondeau DP, Hura GL, Hitomi C, Berglund JA, Getzoff ED, Cary SC, Tainer JA. 2009. Superoxide dismutase from the eukaryotic thermophile *Alvinella pompejana*: structures, stability, mechanism, and insights into amyotrophic lateral sclerosis. *J Mol Biol.* 385(5):1534–1555.
- Spagnolo L, Törö I, D'Orazio M, O'Neill P, Pedersen JZ, Carugo O, Rotilio G, Battistoni A, Djinnovic-Carugo K. 2004. Unique features of the SODC-encoded superoxide dismutase from mycobacterium tuberculosis, a fully functional copper-containing enzyme lacking zinc in the active site. *J Biol Chem.* 279(32):33447–33455.
- Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, ER, van Schleper, C, Guy L, Ettema TJG. 2015. Complex archaea bridge the gap between prokaryotes and eukaryotes. *Nature* 521(7551):173–179.
- Stecher G, Tamura K, Kumar S. 2020. Molecular evolutionary genetics analysis (MEGA) for macOS. *Mol Biol Evol.* 37(4):1237–1239.
- Steinman HM. 1987. Bacteriocuprein superoxide dismutase of *Photobacterium leiognathi*. isolation and sequence of the gene and evidence for a precursor form. *J Biol Chem.* 262(4):1882–1887.
- Sturtz LA, Diekert K, Jensen LT, Lill R, Culotta VC. 2001. A fraction of yeast Cu,Zn-superoxide dismutase and its metallochaperone, CCS, localize to the intermembrane space of mitochondria. A physiological role for SOD1 in guarding against mitochondrial oxidative damage. *J Biol Chem.* 276(41):38084–38089.
- Subramaniam JR, Lyons WE, Liu J, Bartnikas TB, Rothstein J, Price DL, Cleveland DW, Gitlin JD, Wong PC. 2002. Mutant SOD1 causes motor neuron disease independent of copper chaperone-mediated copper loading. *Nat Neurosci.* 5(4):301–307.
- Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol.* 56(4):564–577.
- Vagin A, Teplyakov A. 1997. MOLREP: an automated program for molecular replacement. *J Appl Crystallogr.* 30(6):1022–1025.
- Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25(9):1189–1191.



- Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol.* 18(5):691–699.
- Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, et al. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr.* 67(Pt 4):235–242.
- Wright GSA. 2020. Molecular and pharmacological chaperones for SOD1. *Biochem Soc Trans.* 48(4):1795–1806.
- Wright GSA, Antonyuk SV, Kershaw NM, Strange RW, Samar Hasnain S. 2013. Ligand binding and aggregation of pathogenic SOD1. *Nat Commun.* 4:1758.
- Zerkle AL, House CH, Brantley SL. 2005. Biogeochemical signatures through time as inferred from whole microbial genomes. *Am J Sci.* 305(6–8):467–502.